DOCUMENT RESUME

ED 445 085 TM 031 762

AUTHOR Bielinski, John; Ysseldyke, Jim
TITLE Factors That Lead to Misinterpretation of Trends in the Large-Scale Test Performance of Students with Disabilities: Data-Based Examples.
PUB DATE 2000-04-25
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Academic Achievement; Achievement Tests; *Disabilities; *Educational Trends; Elementary Secondary Education; Regular and Special Education Relationship; *Special Education; State Programs; Student Evaluation; *Test Interpretation; Test Results; *Testing Programs

ABSTRACT

The 1997 reauthorization of the Individuals with Disabilities Act (IDEA) included provisions specifying that states and districts must include students with disabilities in their assessment systems with reasonable accommodations, that states report on the performance of students with disabilities with the same detail and frequency they report on the performance of regular education students. Using a large longitudinal database, this study examined the effect that transitions between regular education and special education across grades had on performance trends for the special education population. The study also examined the effect that changes in exemption rates across grades for students with disabilities had on performance trends. Findings show that the highest achieving special education students left special education and were replaced by the lowest performing regular education students. The result was a substantial increase in the performance gap between regular education and special education students across grades. Furthermore, the reduction in exemption rates from testing at higher grades added to the size of the gap. (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: TM

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. EC

ED 445 085

Factors that Lead to Misinterpretation of Trends In the Large-Scale Test Performance of Students with Disabilities: Data-Based Examples

John Bielinski & Jim Ysseldyke

National Center on Educational Outcomes

University of Minnesota

TM031762

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

J. Bielinski

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract

The 1997 re-authorization of the Individuals with Disabilities Act (IDEA) included provisions specifying that states and districts must include students with disabilities in their assessment systems with reasonable accommodations, that states report on the performance of students with disabilities with the same detail and frequency they report on the performance of regular education students. Using are large longitudinal database, this study examined the affect that transitions between regular education and special education across grades had on performance trends for the special education population. This study also examined the affect that changes in exemption rates across grades for students with disabilities had on performance trends. The study found that the highest achieving special education students left special education and they were replaced by the lowest performing regular education students. The result was a substantial increase in the performance gap between regular education and special education students across grades. Furthermore, the reduction in exemption rates from testing at higher grades added to the size of the gap.

Factors that Lead to Misinterpretation of Trends in the Large-Scale Test Performance Of Students with Disabilities: Data-Based Examples

Educators, parents, policy makers and citizens want and need to know the extent to which all students, including students with disabilities, are profiting from their educational programs. Recent school reforms are resulting in greater attention being paid to the educational results for students with disabilities. The 1997 re-authorization of the Individuals with Disabilities Act (IDEA) included provisions specifying that states and districts must include students with disabilities in their assessment systems with reasonable accommodations, and beginning July 1998 had to report on the performance of students with disabilities on alternate assessments. States are to disaggregate data on the performance of students with disabilities with the same detail and frequency they report on the performance of regular education students. IDEA provisions have resulted in more publicly available data of educational results for students with disabilities than has ever been available in the past. A recent report by the National Center on Educational Outcomes found that 18 states publicly report achievement results for students in special education (Thurlow, Nelson, Teelucksingh, & Ysseldyke, in press). Each year more states are coming on line with results.

Because the results fill an important void, and because they are readily available in state reports, the public, the media, policy-makers, and educational researchers will be inclined to use the results to interpret performance levels and performance trends for students with disabilities. Purportedly the main reason for disaggregation is to obtain some indication of how well education is working for students with disabilities and then, over time to use such data to make improvements in the delivery system. A central concern is whether students with disabilities are

benefiting from education to the same degree as their non-disabled peers. To do this, however, requires clear and consistent groups over time. Special education status complicates the reporting because group membership may change. This report documents several factors that, when not taken into consideration will lead to misinterpretation of trends in the performance of students with disabilities across grades.

Achievement: Special Education vs Regular Education

Currently, large-scale performance trend data for students with disabilities is rarely available. The few results that have been reported paint a grim picture of the performance of students with disabilities. The achievement gap between students with disabilities and their non-disabled peers grew steadily across grades. National Center on Educational Outcomes Twelve summarized passing rates for students with disabilities and students without disabilities across several states. Although the criteria for passing varied across states, in every state passing rates were lower for students with disabilities. The difference in passing rates between the two groups ranged from 23% to 47%, with an average difference of 37%. Regular education students were 2.5 times more likely to meet minimum state requirements in 8th grade than were the students with disabilities (Ysseldyke, Thurlow, Langenfeld, Nelson, Teelucksingh, Seyfarth, 1998). Similar differences were shown in mathematics.

Another report tracked performance trends across grades in Kentucky, reporting the achievement gap as effect sizes. The gap between the disabled and non-disabled students on the 1996 reading test, was less than .25 in 4th grade, almost .70 in 8th grade reading test, and over .80 in 11th grade. The size of the achievement gap quadrupled from 4th grade to 11th grade! Assuming a normal test score distribution, the average score of 4th graders with disabilities fell at

the 40th percentile of the non-disabled group test score distribution, whereas the average score of 11th graders with disabilities fell at just the 21st percentile of the non-disabled group test score distribution.

The question we consider in this report is, "Can disaggregated test results for students with disabilities be used to make valid inferences as to whether students with disabilities are benefiting from education to the same degree as their non-disabled peers?" The expectation would be that achievement growth rates for students with disabilities should be similar to the non-disabled students. Evidence of this would be revealed as a steady performance gap across grades. That is, the achievement gap between students with disabilities and non-disabled students should be the same at 11th grade, as it is at 3rd grade. Findings described above suggest that is not the case. If these trends across grades validly reflect the degree to which students are benefiting from education, then the appropriate conclusion would be that students with disabilities are not benefiting from education to the same degree as their non-disabled peers.

In this report we demonstrate how certain factors influence performance trends for students with disabilities as a group differently then for the group of non-disabled students, and that disregarding these factors will lead to invalid inferences about achievement growth for students with disabilities. Two factors are examined. One factor is that disability status is dynamic. It is not uncommon for a student to be labeled as disabled for reporting purposes one year, but labeled as non-disabled the following year. A classical example is speech/language disability. Prior to 5th grade, this is one of the most commonly used disability categories, but it is rarely found in students after 5th grade. Disability classification is often tied to achievement such that low achieving students are more likely to get a disability label. Students are also declassified, but little has been reported as to how declassification is tied to achievement. The

situation is further complicated when states record performance by special education status and not disability status because special education status is even less stable across grades.

The other factor influencing performance trends for students with disabilities is testing rates. Several reports shown that testing rates among students with disabilities are higher in secondary school than in primary school (Trimble, 1998). It is likely that disabled students tested in primary school are a more select group (with respect to achievement) than are the disabled students tested in secondary school, thus resulting in a larger achievement gap.

Methods for Reporting Trends

Any interpretation of performance trends should always consider the manner in which trends were recorded. In order to talk about trends, one must have longitudinal data. There are two common data collection practices for generating longitudinal performance trends: one is to gather data from a cross-section of examinees across several grade levels, and is heretofore called the cross-sectional method; the other is track a cohort over several years, and is heretofore called the cohort-method. The principal advantage of the cross-sectional method is that performance trends can be recorded using single year's test results. Most states report results from tests administered at different grades. Trends in the performance of various groups can be described by comparing the size of the gap at different grades. The results available in most state reports are based on cross-sectional data collection. A disadvantage with this method is that the gap is calculated on different students at each grade. If the characteristics of a particular group (e.g. students with disabilities) change substantially across grades, then one is really comparing apples (e.g. 4th graders) to oranges (e.g. 8th graders); they're different students.

With the cohort method, the gap at a higher grade (e.g. 8th grade) is calculated on the scores of students who were used to calculate the gap in an earlier grade (e.g. 4th grade). Barring high dropout rates (usually not an issue at 8th grade) or some dramatic shift in the demographics in an entire state, of a particular group of students (highly unlikely over a four year period), changes in the gap will reflect real differences in achievement growth rates for each group. The principal shortcoming of this method is that states rarely stay with a particular testing program long enough to generate longitudinal data on cohorts.

The cohort method of generating longitudinal data may be further divided into approaches. One approach was discussed above; the gap at 8th grade was compared to the gap at 7th grade from the prior year's test scores, and so on back to 4th grade, using data collected 4 years prior. We refer to this method as the cohort-dynamic method, because group composition may change from one year to the next. The other method would require that the same students be repeatedly tested at each grade. We refer to this method as cohort-static.

One challenge to reporting performance trends is that meaningful comparisons across different tests (tests designed for different grades) require a common metric or scale. This is accomplished through a method called vertical scaling. The method contains many unresolved problems. Because of the problems inherent in vertical scaling, this report examines trends in the performance gap between groups of interest, e.g. regular education students versus special education students. A common metric for reporting group differences is called the effect size. An effect size expresses the discrepancy between two means relative to the standard deviation of the test scores. Equation 1, below shows how the effect size is calculated.

$$ES = \frac{(\bar{X}_{Reference} - \bar{X}_{Focal})}{S_{Reference}} \qquad (1)$$

The numerator represents the mean achievement difference between the reference group (regular education) and the focal group (special education). The denominator represents the standard deviation of the test scores for the reference group. An effect size of ES = .50 would indicate that the mean score for the reference group was one-half of a standard deviation unit above the mean score of the focal group. Assuming a normal score distribution, one could say that the mean score for focal group fell at the 30th percentile of the reference group score distribution.

Procedures

The first objective of this study was to report trends in the performance gap between regular education and special education students. Three methods were used to define groups and report trends. The three methods are defined as:

(1) **Cross-sectional** – As is typical for cross-sectional data, the performance gap between special ed. and regular ed. students is reported using data from a single administration of the TAAS. The 1995 TAAS scores were chosen because that year falls in the middle of the range of years on which we had data. The gap within each grade, 4th through 8th, was computed. In this analysis, one group represents the students receiving special education services in that grade, and the other group represents kids not receiving special education services (i.e. regular ed group).

(2) **Cohort static** – Group membership (i.e. special ed. or regular ed.) was determined by each student's status in 1993, and it remained static across years. Test performance for each group was tracked over five years, and the gap in each year computed.

(3) **Cohort dynamic** – As with the cohort static method, the sample was comprised of all 4th graders tested in 1993, and test performance was tracked for five consecutive years.

Group membership (special education vs. regular education), however, was redefined each year. Therefore, in each year, the students receiving special education services in that grade were always compared to the students NOT receiving special education services (i.e. regular ed. students). Because special education status may vary over time for each student, group membership was dynamic.

Trend-lines were graphed, and visual analysis was used to interpret the findings. It was unnecessary to conduct inferential analysis because the entire population of students was used, therefore, any differences represent population differences and are not affected by sampling error.

The next phase of analysis demonstrates how the trend lines under the various methods are affected by changes in special education status. Four groups of students were defined: (1) students who remained in regular education from one year to the next; (2) students who moved from regular education to special education; (3) students who moved from special education to regular education; and (4) students who remained in special education from one year to the next. Membership in these four groups was determined solely by status in adjacent years. Test performance was based upon the first of the two adjacent years.

The last phase of analysis compared test performance for students tested in one year but not the next, with those tested in consecutive years. The comparison was made for both special education and regular education students. If exemption from testing is a function of performance, wherein lower performing students are more likely to be exempt, than students tested in consecutive years will outperform those tested in one year but not the next.

Results

Transitions

Table 1 shows the rate at which students transition between regular education and special education. Two group were defined: those that moved from regular education to special education the following year (labeled Enter); and those that moved from special education to regular education (labeled Left). There were 25,799 students receiving special education services in 4th grade in 1993. Among those, 3316 or 12.9% did not receive services in 1994, while 4626 students (17.9%) that were not receiving special education services in 1993, received special education services in 1994. No students who may have entered the state after 1993 were included.

The rate at which students were added to special education was higher than the rate of leaving special education after every grade but 7th grade. The rate of leaving special education or entering special education decreased from one grade to the next. The rate of entering special education dropped from 17% after 4th grade to 8.0% after 7th grade. The rate at which students left special education decreased from 12.9% after 4th grade to 9.6% after 7th grade.

Table 1

The percent of the total special education population by sex entering and leaving special education each year.

	Grade							
	4		5		6		7	
	Left	Enter	Left	Enter	Left	Enter	Left	Enter
Total	12.9	17.0	11.1	13.0	9.3	10.6	9.6	8.0
Male	12.1	16.2	10.4	12.8	8.5	10.0	8.8	7.6
Female	14.5	18.8	12.3	13.3	10.9	11.6	11.2	8.3

In any given year, roughly 80% of the special education population remained unchanged. If one were to assume that a student's status changed only once, then the transition rates reported here would mean that the special education population would completely turn over in five years! However, as it is, two-thirds of the original special education population remained in special education across all five years. Not only did kids leave special education, but, some of them returned to special education. For instance, 541 (or 16%) of the kids who left special education after 4th grade, returned after 5th grade. In total, over 39,000 kids changed status at least once over the five years, of them, over 6,100 changed status at least twice. The transient nature of special education status begs the question: "To whom is the performance of this transient group attributed, special education or general education?"

Performance Trends

Figures 1 and 2 show the trend in the achievement gap between the reference group and the focal group across five consecutive years. The precise characteristics of the groups depends upon which method was used (i.e. cross-section, cohort-dynamic, & cohort-static). For the cross-sectional and cohort-dynamic comparisons, the reference group is entirely of students in regular education, and the focal group is composed entirely of students in special education. For the cohort-static comparison, the reference group is composed of all students who were in regular education in 1993, regardless of their status in subsequent years. Likewise, the focal group is composed of all students who were in special education in 1993 regardless of their status in subsequent years. The x-axis represents the year the test was given, and the y-axis represents the achievement gap between special and regular education students defined by the effect size. Although not explicitly shown, the horizontal line equal to zero represents the achievement of the regular education students[1]. Effect sizes were computed such that negative values indicate that the focal group (i.e. special education) mean was below that of the reference group (regular education) mean. Only students that had a valid test score in all five years were included in this analysis.

Because of the way in which group membership was defined, the cohort static and the cohort dynamic trend lines were consisted of identical groups in 4th grade. Under all three methods of tracking grade trends, the mean achievement score for students in special education fell substantially below the regular education mean achievement score. However, after that grade the trend lines diverge. Under the *cohort-static* method (where group membership

[1] In the cohort-static comparison, the reference group was defined as students who were NOT receiving special education services in 4th grade.

remained the same across grades), the reference group gained some ground. The effect size for the reading test score decreased slightly from -.48 in 4th grade to -.42 in 8th grade. This trend contrasts with the increasing achievement gap under the *cohort-dynamic* method, in which the effect size dropped precipitously from -.48 in 4th grade to -.93 in 8th grade. Translated in terms of percentile ranks, the mean test score for the special education students in 4th grade fell at the 31st percentile of the regular education score distribution. By 8th grade, under the cohort-static method the mean for the focal group (most special education students) fell at the 28th percentile of the regular education score distribution, while under the cohort-dynamic method, the mean for the focal group (students in special education in 8th grade) dropped to the 18th percentile of the regular education score distribution. The *cross-sectional* trend line mirrored the *cohort-dynamic* trend line. Among 4th graders taking the test in 1995, the effect size between regular education and special education students was -.64, and it grew to –1.16 by 8th grade. The gap increased by more than .50 standard deviation units.

The findings in math achievement were similar. The achievement gap from 4th to 8th grade increased .52 standard deviation units under the *cross-sectional* method, .56 standard deviation units using the *cohort dynamic* method, but it *decreased* .04 standard deviation units under the *cohort static* method.

Figure 1

The READING test score effect size between regular education and special education students.

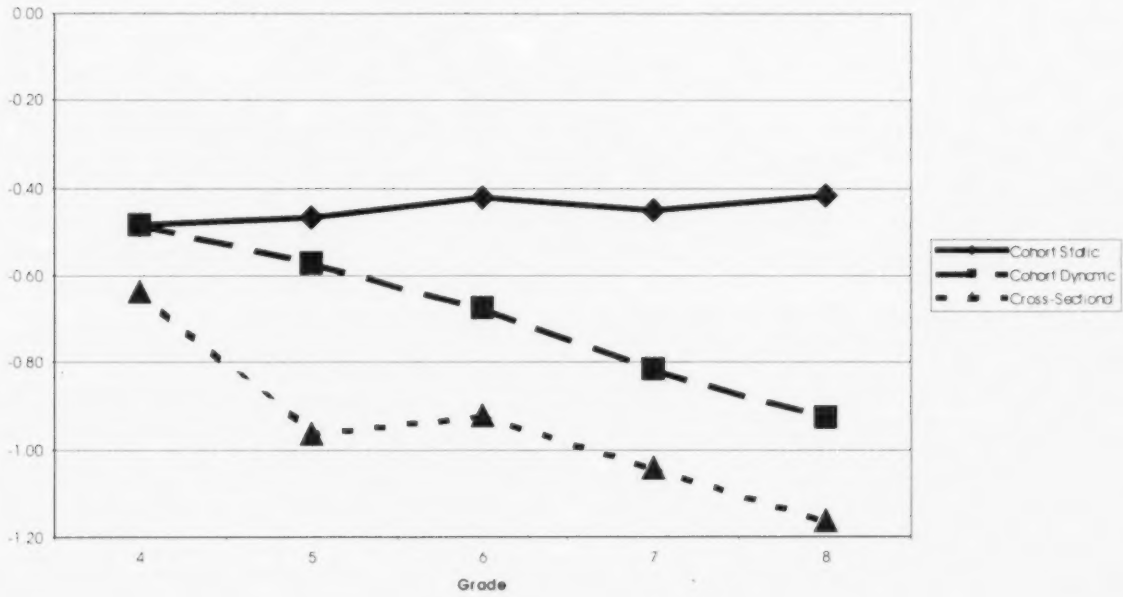
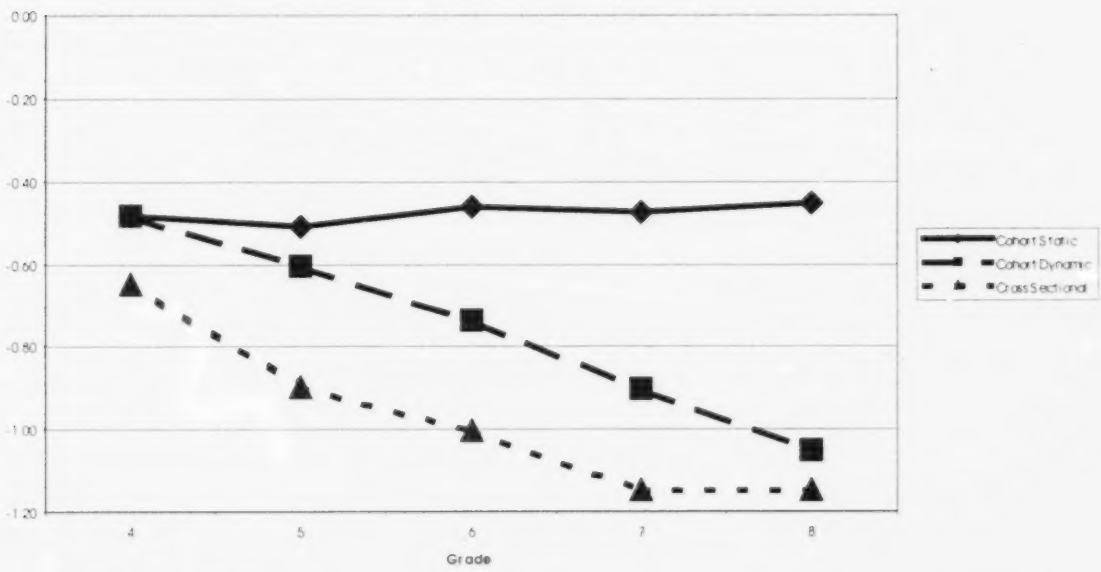


Figure 2
The MATH test score effect size between regular education and special education students.



Achievement & Transitions

We speculated that the reason the trend lines differ so dramatically is that the characteristics of students in special education change dramatically across grades, and that this change is tied to student achievement. To demonstrate how transitions between regular education and special education are tied to achievement we examined the relationship between achievement and movement into and out of special education. Each year, the mean score for students representing four groups was determined. The groups were defined by special education and regular education in two consecutive years. Group 1 (labeled G1) represented students who remained in regular education, Group 2 (labeled S1) were those who moved out of regular education and into special education, Group 3 (labeled G2) were those who moved out of special education and into regular education, and Group 4 (labeled S2) were those who remained in special education both years. The mean reading achievement score (TLI score) for each group at each grade is shown in Figure 3. Group membership is based upon each student's status in the year following the test score. For instance, the mean reading score for 4th grade students who remained in regular education in 5th grade was 74.5; whereas the score for the regular education students who moved into special education was 58.9.

There are two particularly important results. First, in each year, the average reading test scores for the students leaving special education was much higher than it was for those who remained in special education. On the reading test, the mean difference between these two groups ranged from between 6 and 9 points, which translates to between one-third and one-half of a standard deviation unit. On the math test, the students leaving special education outperformed those remaining in special education by at roughly 9 points, or about one-half of a

standard deviation unit. The other important feature of the results is that the students leaving special education were replaced by a much lower achieving group of regular education students. In fact, the group entering special education had even lower achievement than the group remaining in special education in two of the four years. On the reading test, the mean difference between those leaving special education and those entering special education ranged from 7 points in 6th grade to nearly 12 points in 4th grade. On the math test, the difference ranged from 7 points in 6th and 7th grade to more than 11 points in 5th grade. In standard deviation units, these results indicate that students leaving special education outperformed those entering special education by as much as .75 standard deviation units.

Figure 3

Mean test performance in reading for kids whose status in the following year either remained as general ed. (G), changed from general ed. to special ed (S), changed from special ed to general ed (G), or remained as special ed. (S).

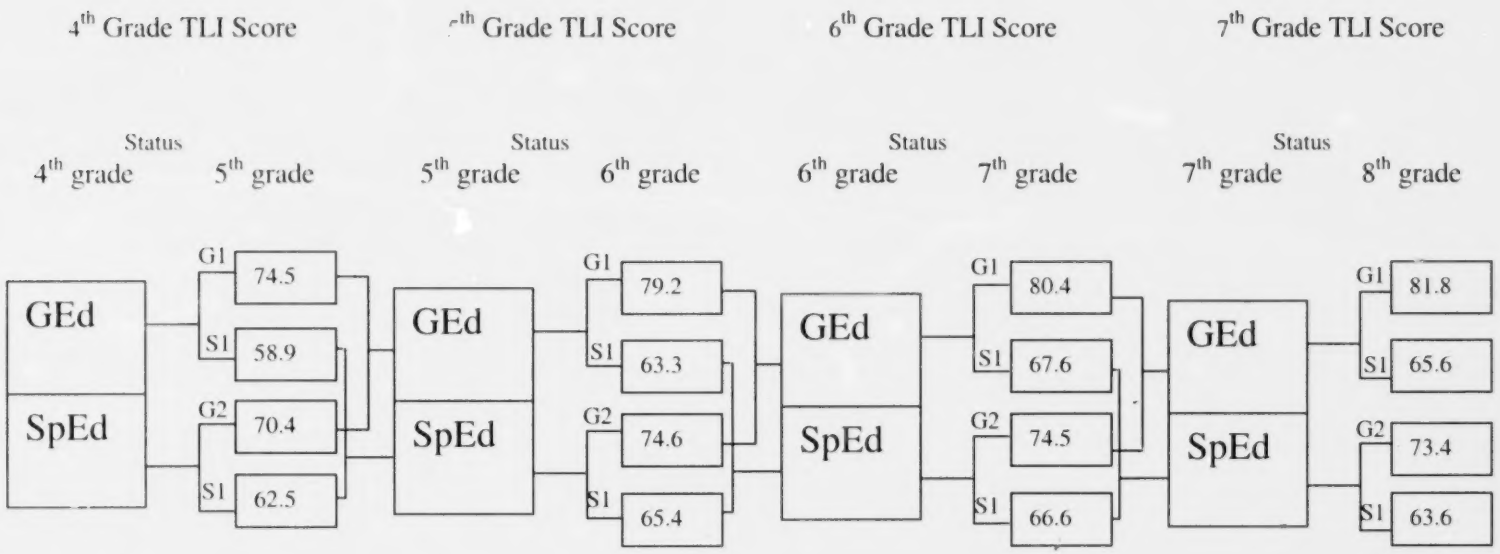
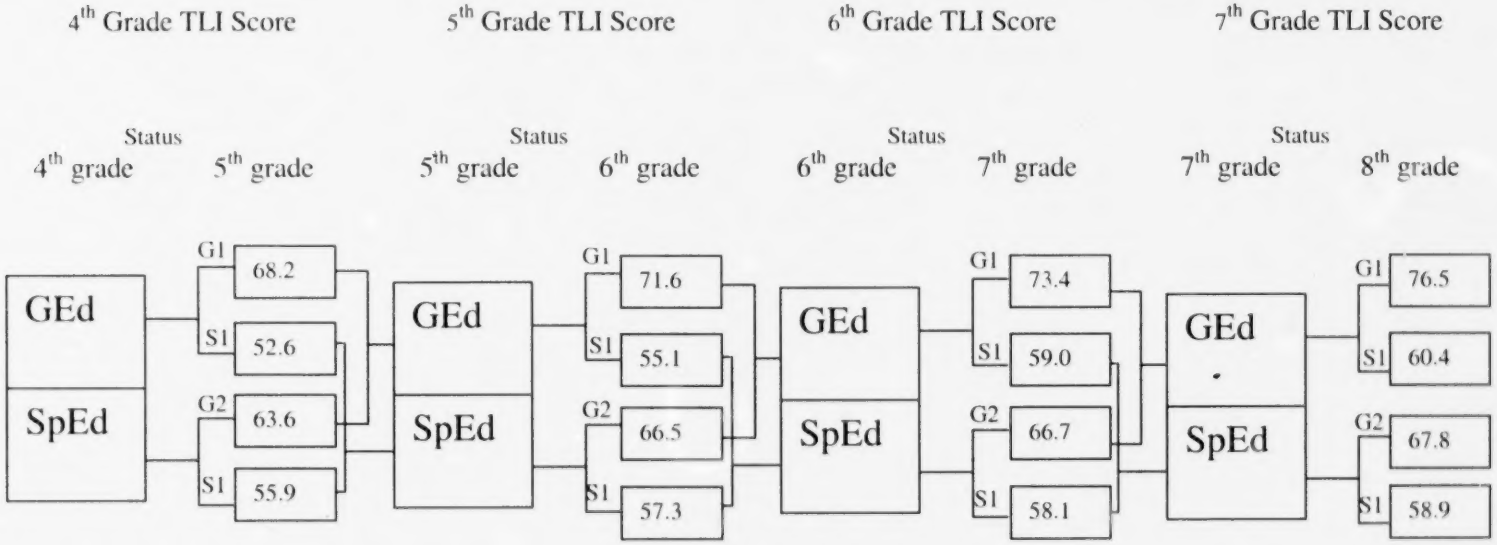


Figure 4

Mean test performance in mathematics for kids whose status in the following year either remained as general ed. (G), changed from general ed. to special ed (S), changed from special ed to general ed (G), or remained as special ed. (S).



If decisions about whether a student should receive special education services are tied to achievement, such that the services are provided only to those students whose achievement falls well below their regular education peers, then one may predict that there would be uncertainty about what to do with students whose achievement is not well below their general education peers. To test this assertion, we calculated the mean achievement score for students who moved between special education and regular education more than once, and compared it to scores for students who moved only once, and to those whose status remained unchanged all five years. We speculated that mean achievement levels for students who moved between special education and regular education more than once, would fall between those who moved once from regular education to special education, and those who moved once from special education to regular education. Five groups of students were defined. The two stable groups, expected to define the achievement extremes, were those that remained in regular education all five years and those that remained in special education all five years (i.e. grades 4 through 8). Two groups were defined by students whose made exactly one transition; either from special education to regular education or from regular education to special education. The final group was comprised of students whose status changed at least twice.

Figures 5 and 6 show the mean test performance at each grade for the five groups. For those who had only one status change, the change could have occurred in any grade 4 through 7. As expected the achievement of students who remained in regular education was the highest and those who remained in special education was the lowest. The mean achievement of students who transitioned back and forth between special education and general education fell between the achievement of those who moved once from special education to general education and those who moved once from general education to special education. In fact, in both reading and math

the "back-and-forth' group's mean test performance fell almost exactly between the performance of those who remained in special education throughout and those who remained in general education throughout.

Figure 5

Mean scaled score on the reading test for mutually exclusive groups defined by general education/special education status.

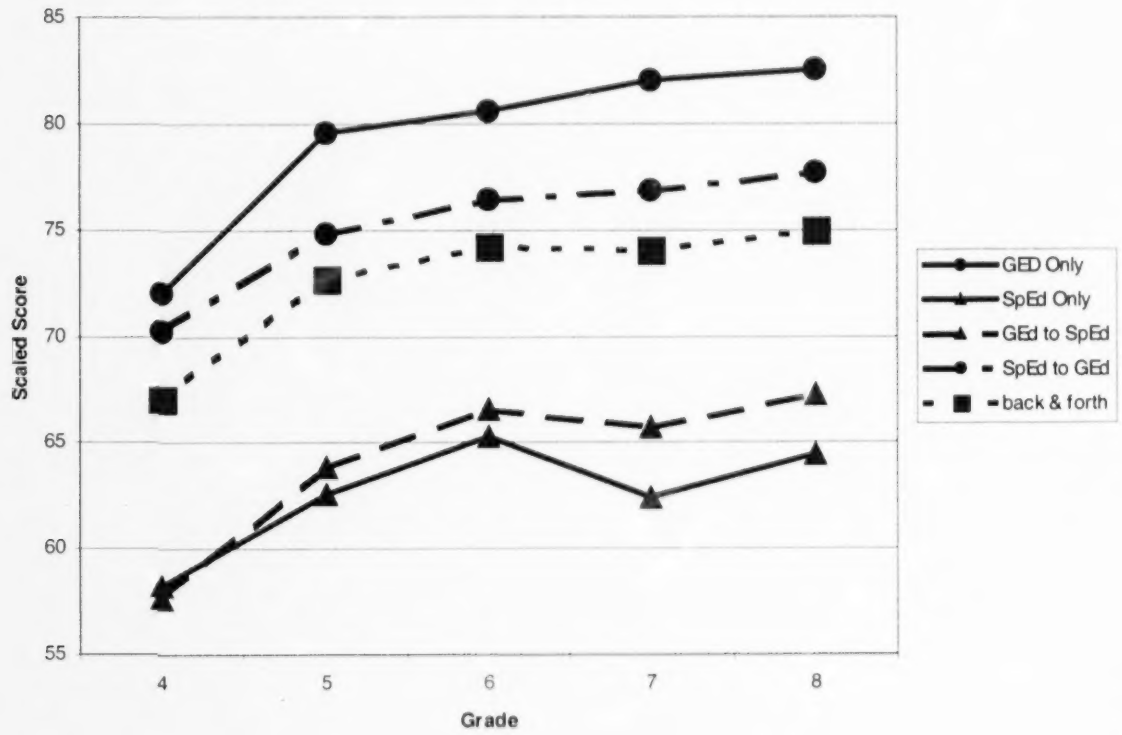
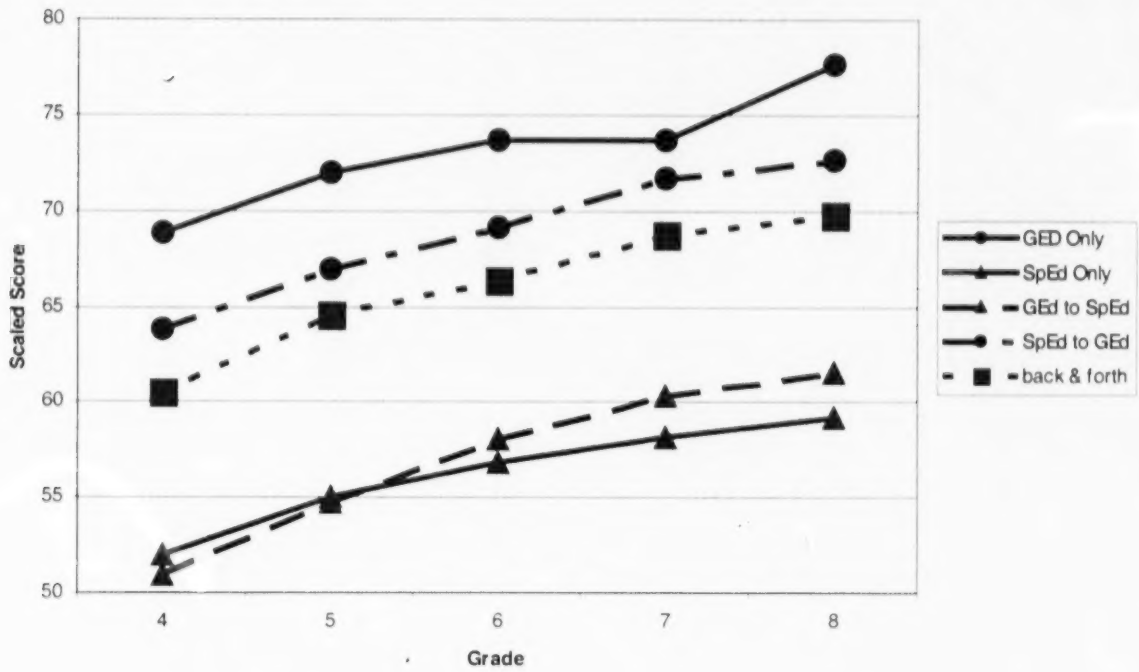


Figure 6

Mean scaled score on the math test for mutually exclusive groups defined by general education/special education status.



Impact of Exemption

Table 2 shows the testing rates by special education status. The percent of the regular education students in each grade that were tested increased from roughly 93% in 4th grade to nearly 97% by 8th grade. The increase in testing rates was more dramatic among students in special education. In 4th grade, only 41.9% of the special education students took the reading test and 47.2% took the math test, by 8th grade the testing rate increased to 61.1% in reading and 60.7% in math. The most notable jump in the testing rate for special education students occurred between 6th and 7th grade where the rate jumped from roughly 45% to nearly 60%.

Table 2. The percent of students with each group that were tested at each grade.

Grade	Percent Tested			
	Reading		Math	
	GEd	SpEd	GEd	SpEd
4	92.9	41.9	93.1	47.2
5	93.6	42.4	93.8	45.6
6	96.2	45.0	96.1	46.6
7	97.4	59.2	97.2	59.5
8	96.9	61.1	96.9	60.7

In order to demonstrate the impact of the increased testing rates, we tracked the achievement gap between students receiving special education services in 4th grade and those not receiving services in 4th grade under two scenarios. Under one scenario, the achievement gap was computed using only students that had a valid test score all five years, therefore, the trend line is the same as the cohort-static trend line in Figures 1 and 2. Under the other scenario, all students with a valid test score in any given year were used to compute the achievement gap. Although students were assigned to a group based upon their status in 4th grade, many of the 4th graders were not tested, therefore, the achievement gap includes different students subsequent grades. The trend lines are shown in Figures 7 and 8. The solid line in each figure is based on students with a valid test score all five years making it identical to the cohort-static trend-line in Figures 1 and 2. The dashed line represents the trend-line when all students tested in a given year are used to compute the achievement gap. The dashed line drops precipitously with increasing testing rates. The achievement gap, represented as an effect size, increased from -.60

in 4th grade to -.94 in 7th grade on the reading test, and from -.65 to -.98 on the math test. These results suggest that the students from the special education group that were exempt from testing in the earlier grades were lower achievers compared to those who were tested every year.

Figure 7

The effect size on the READING test across grades between students who were receiving special education services in 4th grade compared to those who were not, when (a) only students with a valid test score across all five years were used; and (b) when all students with a valid test score each year were used.

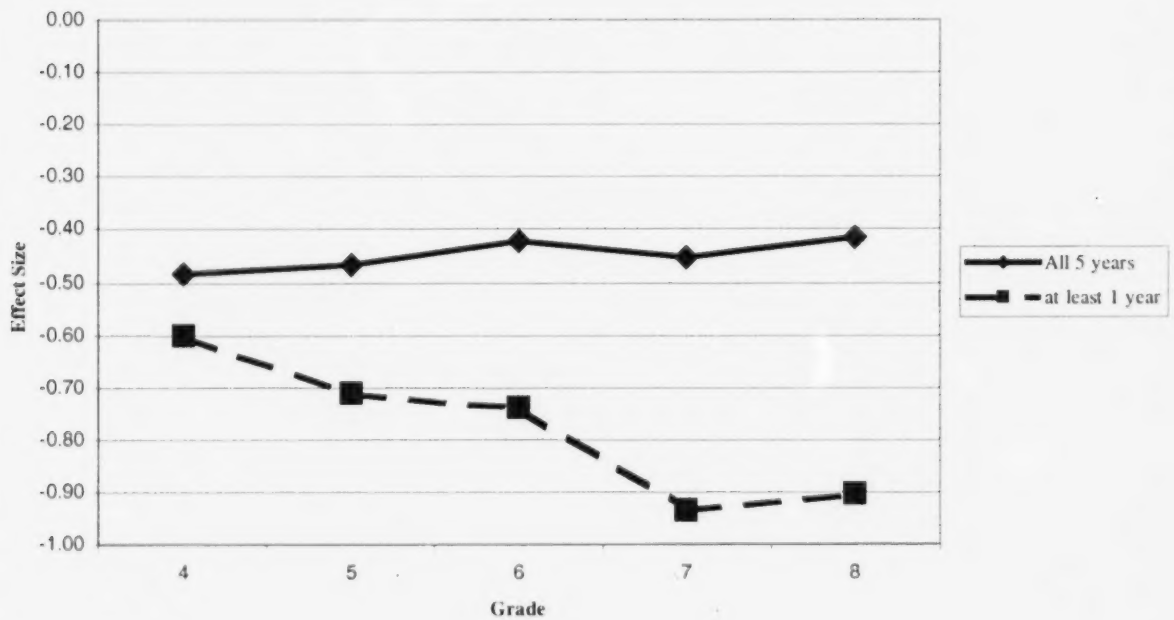
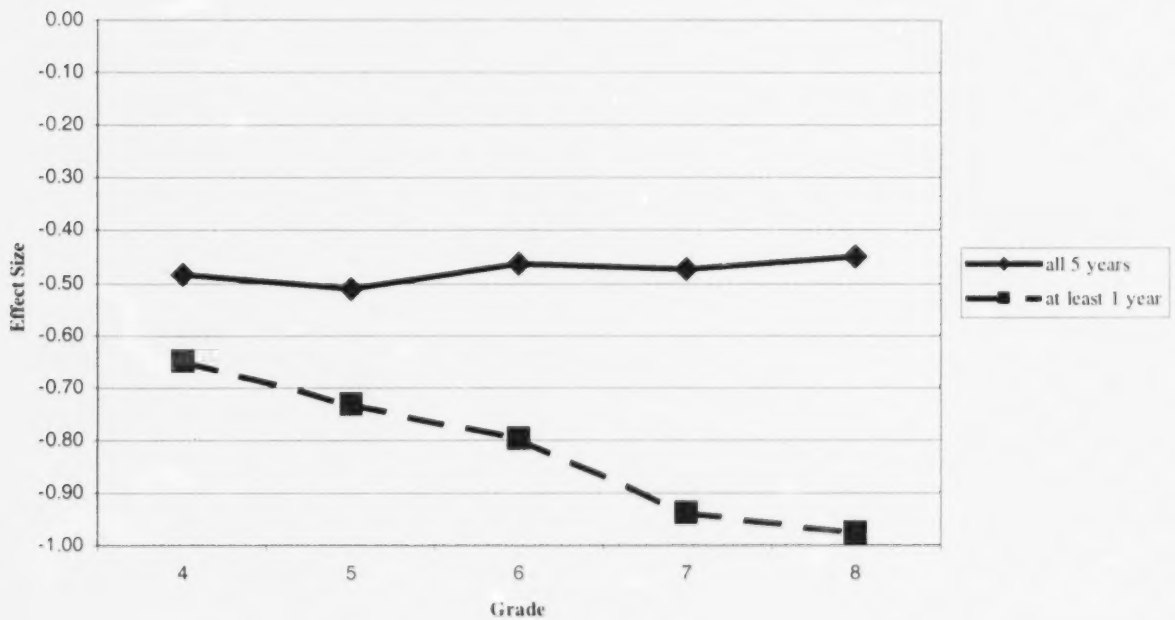


Figure 8

The effect size on the MATH test across grades between students who were receiving special education services in 4th grade compared to those who were not, when (a) only students with a valid test score across all five years were used; and (b) when all students with a valid test score each year were used.



The results in Table 3 buttress this argument. Table 3 compared achievement between students tested in two consecutive years with those tested one-year, and not the next. At each grade level, the students that had been tested in the previous grade, but not the current grade, had a substantially lower achievement than those tested at each grade level. Among students in special education the achievement a high of 13.8 points in 4th grade to 9.5 points in 7th grade. The pattern among regular education students was similar. In 4th grade, the mean reading score of those students exempt from testing in 5th grade was 16.1 points lower than the score of those

students tested both years. By 7th grade, the difference dropped to 8.7 points. Table 3 also shows the impact those students not tested both years had on the overall mean for each group. Among regular education students, the mean TLI score on the 4th grade reading test was 74.8 for those students tested in consecutive years and it was 74.2 when students tested that year but not the next were included. In other words, the students exempt from testing the following year brought down this group's score only .6 points. The influence on mean test performance of exempt students was greater among special education group. The mean reading TLI score was 66.5 for the special education students tested in consecutive years and it was 64.5 with the students tested that year but not the next included. Across each grade on both the reading and the math test, the influence on mean performance of students not tested in consecutive years was greater for the special education group than it was for the regular education group.

Table 3

The mean TAAS reading score by disability status for students tested in two consecutive years and those tested one year but not the next.

		Mean Test Score			
Grade Tested		Tested two consecutive years	Tested one year but not the next	Difference	Combined
4	GEd	74.8	58.7	16.1	74.2
	SpEd	66.5	52.7	13.8	64.5
5	GEd	79.4	66.6	12.8	79.0
	SpEd	69.2	57.2	12.0	67.4
6	GEd	80.4	70.5	9.9	80.2
	SpEd	68.7	59.2	9.5	67.8
7	GEd	81.9	73.2	8.7	81.6
	SpEd	66.2	54.9	11.3	64.9

The reason that the influence on mean test performance was greater in the special education group, even though the mean difference was similar across groups was that the exempt students represented a much larger proportion of the special education group than they did in the regular education group. Table 4 shows the number and percent of kids tested in one year and not tested the following year by special education status. Among the special education students, the percent ranged from 9.0% in 6th grade to 14.7% in 4th grade. Among the general education students, the percent ranged from 2.2% in 6th grade to 4.1% in 4th grade. Special education

students that were tested one year were three times more likely not to be tested the following year than kids in general education.

Table 4

Number and percent of kids tested one year and not tested the next by special education status.

	Group			
	Special Ed.		General Ed.	
Grade Tested	N	%	N	%
4	1587	14.7	7367	4.1
5	1667	14.5	4280	2.4
6	1123	9.0	3969	2.2
7	1990	11.9	5459	3.0

Discussion

The purpose of this study was to ascertain how transitions between regular education and special education, and increasing testing rates affect the size of the achievement gap between regular education and special education students. Our analysis determined that roughly 20% of the special education population in a given year, had experienced a status change the prior year. This finding is not exclusive to this state. The Twentieth Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act reported that, among adolescents 14 years old and older in special education, over 63,000 or roughly 3 to 4% of the population returned to regular education in the 1995-96 school year (US Dept of Education,

Office of Special Education Table AD1, 1998). The report did not include declassification data on students under the age of 14. At least two studies reported declassification rates for elementary grade students. Walker, Singer, Palfrey, Orza, Wenger, & Butler, (1988) reported a two-year follow study that tracked 1,184 elementary students who were in special education. Of those remaining in the district at the end of two years, 17% were no longer receiving special education services. In a separate study Carlson and Parshall (1996) used data collected by the Michigan Department of Education to track special education declassification over a five year period. Results were based on child count data used to satisfy federal reporting requirements. They found that each year 7% of students receiving special education services in Michigan schools return to full-time general education. They found that about 4% of students who were declassified returned to special education within three years. Unlike the federal data which are constrained to students 14 years old and older, these data included children as young as six years old. The rate of declassification was greatest for students between the ages of eight and 11. Declassification rate dropped steadily after age 10. These data demonstrate that changes in the composition of the students served by special education may be considerable from one grade to the next. Failure to account for these changes could results in misinterpretation of the effectiveness of special education services.

One of the principal question examined in this study was whether transitions between special education and regular education affect the magnitude of the achievement gap across grades. We compared trend-lines under three conditions of reporting longitudinal data: cross-sectional; cohort-dynamic; and cohort-static. The cohort-static method illustrated what happens to the achievement gap across grades when the composition of both the regular education and special education groups are held constant. The cohort-dynamic method illustrated the situation

when group composition changes result only from the transitions between regular education and special education, but not from changes due to increasing testing rates. The cross-sectional method illustrates the scenario when group composition changes result from both transitions between regular education and special education and increasing testing rates across grades.

When the same two groups of students were tracked across grades (cohort-static), students in the special education group actually closed the gap slightly! On the other hand, when trend-lines were reported such that transitions between regular education and special education were not controlled for (cohort-dynamic), the gap increased dramatically. To understand why this occurred, we examined achievement levels for students leaving special education and compared them to those entering special education. In each grade, students leaving special education were the highest performing students, with a performance level fairly close to the regular education group. On the other hand, students entering special education were the lowest achieving regular education students. In many instances, this group's achievement was actually lower than that of the students who remained in special education. In other words, over time special education becomes a reservoir for the lowest achieving students and this fact is independent of special education's effectiveness.

The other factor affecting performance trends for students with disabilities examined in this study was exemption rates. First, it was shown that exemption rates were much higher among special education students that regular education students. However, across grades the exemption rates of regular education students remained nearly constant, whereas the exemption of special education students dropped precipitously. The reason is that students are not exempted at random, therefore the higher exemption rates in the early grades results in a more select group of higher achieving special education students being tested. By excluding more of

the low achieving students in the early grades, the gap between regular education students and special education students is artificially minimized.

The problem facing states is that in order to meet reporting requirements outlined in IDEA they may be producing reports that may do more to mislead than inform. Without careful consideration of the factors outlined in this report, transitions between regular education and special education and exemption rates, it would be easy to conclude that special education is ineffective. As student proceed through school, those receiving special education services fall further and further behind their regular education peers. The problem, as illustrated in this report, is that the characteristics of the special education students participating in reporting systems vary dramatically across grades, and, these changes are tied directly to achievement.

One thing States could do to mitigate the problem would be to ensure that all students with disabilities are included in the testing system at every grade level tested. Secondly, states could increase their ability to link their test results from one year to the next such that performance for the same group of students could be tracked over time. In so doing, States could ensure that trends across grades and between groups are always comparing "apples" to "apples." At a minimum, a footnote should be added that indicates that any analysis of trends across grades using only the data provided in a accountability report would result in comparisons between different students at different grades, and may not validly reflect achievement growth of either group.

References

Carlson, E., & Parshall, L. (1996). Academic, social, and behavioral adjustment for students declassified from special education. Exceptional Children, 63, 89-100.

Thurlow, M., Ysseldyke, J., Nelson, J.R., & Teelucksingh, E. (in press). Where's Waldo: ? (Technical Report 23). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Trimble, S. (1998). Performance trends and use of accommodations on a statewide assessment (Maryland/Kentucky Report 3). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

U.S. Department of Education (1998). Twentieth annual report to congress on the implementation of the Individuals with Disabilities Education Act. Washington, D.C.: Author.

Walker, D., Singer, J., Palfrey, J., Orza, M., Wenger, M., & Butler, J. (1988). Who leaves and who stays in special education: A 2-year follow-up study. Exceptional Children, 54, 393-402.

Ysseldyke, J., Thurlow, M., Langenfeld, K., Nelson, J.R., Teelucksingh, E., & Seyfarth, A. (1998). Educational results for students with disabilities: What do the data tell us? (Technical Report 23). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.